Prelimenary results on the optimal control of linear complementarity systems

Alexandre Vieira, Bernard Brogliato, Christophe Prieur

On the optimal control of linear complementarity systems

Alexandre Vieira * **Bernard Brogliato** * **Christophe Prieur** **

* *INRIA Grenoble Rhône-Alpes, Univ. Grenoble-Alpes, 655 avenue de l'Europe, 38334 Saint-Ismier, France, (e-mail: alexandre.vieira@inria.fr - bernard.brogliato@inria.fr)*
** *Univ. Grenoble Alpes, CNRS, GIPSA-lab, F-38000 Grenoble, France (e-mail: christophe.prieur@gipsa-lab.fr)*

Abstract: In this paper, we give some results concerning the quadratic optimal control of linear complementarity systems. We derive first order conditions for this system, that we manage to express as a Mixed Linear Complementarity System. We then use this result to build numerical schemes, which are expressed as Mathematical Problems with Equilibrium Constraints.

Keywords: Linear optimal control, Complementarity problems.

1. INTRODUCTION

We are interested in dynamical control systems having the form:

$$\begin{cases} \dot{x}(t) = Ax(t) + B\lambda^x(t) + Fu(t), \\ y(t) = Cx(t) + d\lambda^x(t) + eu(t), \quad \text{a.e. on } [0,T] \\ x(0) = x_0, \ x(T) \text{ free} \end{cases} \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$, $x(\cdot), B, F \in \mathbb{R}^n$, $C \in \mathbb{R}^{1\times n}$, $d, e \in \mathbb{R}$, $d > 0$, $T > 0$, and $u, \lambda^x : [0,T] \to \mathbb{R}$ are measurable. In order to avoid trivial cases, we assume that $(C, e) \neq (0, 0)$. Furthermore, the trajectory of this system has to comply with the following *complementarity* conditions:

$$0 \leq \lambda^x(t) \perp y(t) \geq 0 \tag{2}$$

meaning that $\lambda^x(t)$,$y(t) \geq 0$ and $y(t)\lambda^x(t) = 0$ for almost all $t \in [0,T]$. The whole system (1)-(2) is called a controlled linear complementarity system (CLCS). This kind of systems, despite their simple look, gives rise to several challenging questions, mainly because conditions (2) introduce non-differentiability at switching points and non-convexity of the set of constraints. It provides a modeling paradigm for many problems, as Nash equilibrium games, hybrid engineering systems (Brogliato (2003)), contact mechanics or electrical circuits (Acary et al. (2011)). Several problems have already been tackled, let us mention observer-based control (Çamlibel et al. (2006), Heemels et al. (2011)) and Zeno behavior (Çamlibel and Schumacher (2001), Pang and Shen (2007), Shen (2014)).

We define now the optimal control problem of finding the trajectories of (1)-(2) minimizing the functionnal:

$$C(u) = \int_0^T f^0(x(t), u(t))dt = \int_0^T (x(t)^\intercal x(t) + u(t)^2)dt. \tag{3}$$

The optimal control of this system is a really challenging question. For instance, the existence of an optimal control still is an open field of research. A well-known theorem, due to Fillipov (see Cesari, 2012, Theorem 9.2i and onwards), requires the convexity of the set

$$\mathcal{U}(x) = \{(u, v) : 0 \leq v \perp Cx + Dv + Eu \geq 0\}$$

which, in this framework, is actually not convex. Another difficulty comes from the fact that the constraints involve both the control and the state. These *mixed constraints* make the analysis even more challenging. For instance, deriving a maximum principle with wide applicability involves the use of non-smooth analysis, even in the case of smooth and/or convex constraints (see e.g. Clarke and De Pinho (2010)). Special cases arise when $y(t) = Cx(t) + d\lambda^x(t)$ (meaning $e = 0$ in (1)). This system can then be seen as an autonomous switching system, where the switching modes are activated when the state reaches some threshold defined by the complementarity conditions (2) (see Georgescu et al. (2012)). The optimal control of such systems has been already studied (see Passenberg et al. (2013) and references therein). Since the control u is also involved in the constraints (2), all these results do not apply. However, in our model, we can get rid of the constraints and compute explicitly λ^x as a function of u and x in a rather simple form. Nonetheless, this expression being non differentiable, we need powerful tools from non-smooth analysis in order to derive a Pontryagin-like maximum principle. The contribution of this paper is that we manage to express these first-order conditions as a Mathematical Problem with Equilibrium Constraints (MPEC) and build on it a numerical scheme. Since the control u was not involved in the constraints in the aforementioned litterature, this approach is new.

This paper is organized as follows: the first section is devoted to the derivation of results leading to a maximum principle for this system. We then recall some results on complete controllability systems, and derive numerical schemes to obtain a numerical solution. We then analyze a simple one-dimensional example, which serves as a benchmark for the numerical schemes. Conclusion ends the paper in Section 7.

2. PRELIMINARY RESULTS

Model (1)-(2) is rewritten as follows:

$$\begin{cases} \dot{x}(t) = Ax(t) + B\lambda^x(t) + Fu(t), \\ 0 \leq \lambda^x(t) \perp Cx(t) + d\lambda^x(t) + eu(t) \geq 0, \quad \text{a.e. on } [0,T] \\ x(0) = x_0,\ x(T) \text{ free} \end{cases} \tag{4}$$

In this case, basic convex analysis proves that $\lambda^x(t) = \frac{1}{d}\Pi_{\mathbb{R}^+}(-Cx - eu)$, where $\Pi_K(x)$ is the projection of x on the set K. Therefore, (4) becomes (where the argument t is omitted for simplification):

$$\begin{aligned} \dot{x} &= f^x(x,u), \\ &= Ax + Fu + \frac{B}{d}\Pi_{\mathbb{R}^+}(-Cx - eu), \\ &= \begin{cases} Ax + Fu & \text{if } Cx + eu \geq 0, \\ \left(A - \dfrac{BC}{d}\right)x + \left(F - \dfrac{e}{d}B\right)u & \text{otherwise,} \end{cases} \end{aligned} \tag{4'}$$

with obvious definition for f^x. In (3), the function under the integral sign is smooth, so its subdifferential only contains the classical derivative:

$$\partial_x f^0(x,u) = \{2x^\intercal\}.$$

The vector field of the dynamical control system is non smooth, but we can compute its subdifferential as:

$$\partial_x f^x(x,u) = \begin{cases} \{A\} & \text{if } \ Cx + eu > 0, \\ \left\{A - \dfrac{BC}{d}\right\} & \text{if } \ Cx + eu < 0, \\ \left[A - \dfrac{BC}{d}, A\right] & \text{if } \ Cx + eu = 0. \end{cases}$$

using the notation:

$$[M_1, M_2] = conv\{M_1, M_2\}$$

for any pair $(M_1,\ M_2)$ of according dimensions matrices and *conv* stands for the convex hull of M_1 and M_2. As stated by Clarke (1976), if the control, and therefore the trajectory, is optimal, then there exist an absolutely continuous function $p : [0,T] \to \mathbb{R}^n$ and a scalar $p^0 \leq 0$ such that $(p, p^0) \neq 0$ and satisfying a Pontryagin-like differential inclusion:

$$\begin{aligned} -\dot{p}^\intercal(t) \in f^p(x,p,u) = p^\intercal(t)\partial_x f^x(x(t),u(t)) \\ + p^0\partial_x f^0(x(t),u(t)). \end{aligned} \tag{5}$$

In our problem, (5) becomes:

$$-\dot{p}^\intercal(t) \in \begin{cases} \{2p^0x^\intercal(t) + p^\intercal(t)A\} & \text{if } Cx + eu > 0, \\ \left\{2p^0x^\intercal(t) + p^\intercal(t)\left(A - \dfrac{BC}{d}\right)\right\} & \\ & \text{if } Cx + eu < 0, \\ \{2p^0x^\intercal(t)\} + p^\intercal(t)\left[A - \dfrac{BC}{d}, A\right] & \\ & \text{if } Cx + eu = 0. \end{cases} \tag{6}$$

Furthermore, the maximum condition on the Hamiltonian holds:

$$\begin{aligned} \langle p(t), f^x(x(t),u(t))\rangle + p^0 f^0(x(t),u(t)) = \\ \max_{v\in\mathbb{R}}\{\langle p(t), f^x(x(t),v)\rangle + p^0 f^0(x(t),v)\}. \end{aligned} \tag{7}$$

Since $x(T)$ is free, there is a terminal condition on p: $p(T) = 0$.

3. (COMPLETE) CONTROLLABILITY CONDITIONS

In order to compute an optimal control, the system should obviously be controllable between the inital and final points. We will only focus on completely controllable systems, relying on Çamlıbel's theorem stating the complete controllability of some LCS. We recall here his result:

Theorem 1. (Çamlibel (2007)). Assume (4) satisfies the following conditions (with $D = d$ and $E = e$):

(1) The matrix D is a P-matrix ; i.e., all its principal minors are positive.
(2) The transfer matrix $E + C(sI - A)^{-1}F$ is invertible as a rational matrix.

Then, (4) is completely controllable if, and only if, the following two conditions hold:

(1) The pair $(A, [F\ B])$ is controllable.
(2) The system of inequalities

$$\eta \geq 0, \tag{8a}$$

$$(\zeta^\intercal\ \ \eta^\intercal)\begin{pmatrix} A - \lambda I & F \\ C & E \end{pmatrix} = 0, \tag{8b}$$

$$(\zeta^\intercal\ \ \eta^\intercal)\begin{pmatrix} B \\ D \end{pmatrix} \leq 0, \tag{8c}$$

admits no solution $\lambda \in \mathbb{R}$ and $0 \neq (\zeta, \eta) \in \mathbb{R}^{n+m}$.

This will be used in Section 6 on a particular case.

4. (4) AND (6) AS A MIXED LCS (MLCS)

We denote $z = \begin{pmatrix} x \\ p \end{pmatrix}$. Let us rewrite (4) and (6) in the compact form:

$$\dot{z} \in \begin{pmatrix} f^x(z,u) \\ f^p(z,u) \end{pmatrix} \tag{9}$$

where the right-hand side is a set-valued function defined by:

- if $Cx + eu > 0$:
$$\begin{pmatrix} f^x(z,u) \\ f^p(z,u) \end{pmatrix} = \left\{\begin{pmatrix} Ax + Fu \\ -2p^0x - A^\intercal p \end{pmatrix}\right\},$$
- if $Cx + eu < 0$:
$$\begin{pmatrix} f^x(z,u) \\ f^p(z,u) \end{pmatrix} = \left\{\begin{pmatrix} \left(A - \dfrac{CB}{d}\right)x + \left(F - \dfrac{e}{d}B\right)u \\ -2p^0x + \left(\dfrac{C^\intercal B^\intercal}{d} - A^\intercal\right)p \end{pmatrix}\right\},$$
- if $Cx + eu = 0$:
$$\begin{pmatrix} f^x(z,u) \\ f^p(z,u) \end{pmatrix} = \begin{pmatrix} \left(A - \dfrac{FC}{e}\right)x \\ -2p^0x + \left[\dfrac{C^\intercal B^\intercal}{d} - A^\intercal, -A^\intercal\right]p \end{pmatrix}.$$

We can recast the differential inclusion (9) in the framework of complementarity systems with linear dynamics:

$$\begin{aligned} \dot{z} = \begin{pmatrix} g^x(z,u) \\ g^p(z,u) \end{pmatrix} = \begin{pmatrix} A & 0 \\ -2p^0I & -A^\intercal \end{pmatrix} z \\ + \begin{pmatrix} B & 0 \\ 0 & \dfrac{C^\intercal B^\intercal}{d} \end{pmatrix}\begin{pmatrix} \lambda^x \\ \lambda_1^p \end{pmatrix} + \begin{pmatrix} F \\ 0 \end{pmatrix}u, \end{aligned} \tag{10}$$

$$0 \le \lambda^x \perp \quad Cx + d\lambda^x + eu \ge 0, \tag{11a}$$
$$0 \le |\lambda_1^{p_j}| \perp \; Cx + eu + |Cx + eu| \ge 0, \tag{11b}$$
$$0 \le |\lambda_2^{p_j}| \perp |Cx + eu| - (Cx + eu) \ge 0 \qquad j = 1...n, \tag{11c}$$
$$|\lambda_1^{p_j}| + |\lambda_2^{p_j}| = |p_j| \qquad j = 1...n, \tag{11d}$$
$$\lambda_1^p + \lambda_2^p = \; p, \tag{11e}$$

where the subscript j denotes the j-th component of a vector, and $\lambda_i^p = (\lambda_i^{p_1}, ..., \lambda_n^{p_n})^\intercal$.

Proposition 2. The right-hand side of (9) is the same as the right-hand side of (10) defined with the complementarity conditions in (11).

Proof. The first line (11a) gives obviously the same right-hand side as in (4), which is $f^x(z,p)$. Then, $f^x(z,u) = g^x(z,u)$. Therefore, we have to check that the other lines in (11) are the same as in (6). To do so, we have to distinguish 3 cases:

- if $Cx + eu > 0$, then from (11b), we deduce that $\lambda_1^p = 0$. It follows that:
$$g^p(z,u) = -2p^0x - A^\intercal p = f^p(z,u).$$
- if $Cx+eu < 0$, then from (11c), $\lambda_p^2 = 0$, and therefore from (11e), $\lambda_1^p = p$. We then have the following equality:
$$g^p(z,u) = -2p^0 + \left(\frac{C^\intercal B^\intercal}{d} - A^\intercal\right) p = f^p(z,u).$$
- if $Cx + eu = 0$, then from (11d), we have that $|\lambda_1^{p_j}| + |\lambda_2^{p_j}| = |p_j|$, so that $\lambda_i^{p_j} \in [-p_j, p_j]$, $i = 1,2$, $j = 1,...n$. But, in order to comply with the last equality (11e), we must have $\lambda_i^{p_j} \in [0,1]p_j$, $i = 1,2$, $j = 1,...,n$, and hence, $\lambda_i^p \in [0,1]p$, $i = 1,2$. This gives us
$$\begin{aligned} g^p(z,u) &= \{-2p^0x - A^\intercal p\} + [0,1]\frac{C^\intercal B^\intercal}{d} p \\ &= \{-2p^0x\} + \left[\frac{C^\intercal B^\intercal}{d} - A^\intercal, -A^\intercal\right] p \\ &= f^p(z,u). \end{aligned}$$

■

We can still have an even more interesting form of (11) by noticing that the function $|\cdot|$ is piecewise linear and so, admits a representation in the form of an LCP. This is the topic of the next lemma:

Lemma 3. Define $\mathbb{1}_n = (1,...,1)^\intercal \in \mathbb{R}^n$. The multipliers λ^x and λ_1^p given by system (11) are equally defined by the following system:

$$\left\{\begin{array}{lll} 0 \le \lambda^x & \perp & Cx + d\lambda^x + eu \ge 0 \\ 0 \le \mu_{x,u} & \perp & \mu_{x,u} - 2(Cx+eu) \ge 0 \\ 0 \le \mu_p & \perp & \mu_p - 2p \ge 0 \\ 0 \le \lambda_1^{abs} & \perp & \mu_{x,u}\mathbb{1}_n \ge 0 \\ 0 \le \lambda_2^{abs} & \perp & (\mu_{x,u} - 2(Cx+eu))\mathbb{1}_n \ge 0 \\ 0 \le \mu_1 & \perp & \mu_1 - 2\lambda_1^p \ge 0 \\ 0 \le \mu_2 & \perp & \mu_2 - 2\lambda_2^p \ge 0 \\ & \lambda_1^{abs} + \lambda_2^{abs} = \mu_p - p & \\ & \lambda_1^p + \lambda_2^p = p & \\ & \mu_1 - \lambda_1^p = \lambda_1^{abs} & \\ & \mu_2 - \lambda_2^p = \lambda_2^{abs} & \end{array}\right. \tag{12}$$

where (12) is a mixed LCP (MLCP).

Proof. First, we need to establish the following simple result: for any scalar x, $|x| = \mu - x$ where μ is given by:
$$0 \le \mu \perp \mu - 2x \ge 0.$$

Indeed, if $x \le 0$, then we must take $\mu = 0$, so that $\mu - x = -x = |x|$. If $x > 0$, then we must take $\mu = 2x$, and $\mu - x = x = |x|$. Let us use (11) to rewrite equivalently the absolute values as:

$$\begin{aligned} &0 \le \mu_{x,u} \perp \mu_{x,u} - 2(Cx+eu) \ge 0, \\ &0 \le \mu_1 \perp \mu_1 - 2\lambda_1^p \ge 0, \\ &0 \le \mu_2 \perp \mu_2 - 2\lambda_2^p \ge 0, \\ &0 \le \mu_p \perp \mu_p - 2p \ge 0, \\ &\lambda_{x,u}^{abs} = \mu_{x,u} - (Cx+eu) = |Cx+eu|, \\ &\lambda_1^{abs} = \mu_1 - \lambda_1^p = |\lambda_1^p|, \\ &\lambda_2^{abs} = \mu_2 - \lambda_2^p = |\lambda_2^p|, \\ &\lambda_p^{abs} = \mu_p - p^\intercal = |p|, \end{aligned}$$

where notation $|x|$ on a vector x is here understood componentwise, e.g. $|x| = (|x_1|, ..., |x_n|)^\intercal$.
Therefore, (11) becomes:

$$\left\{\begin{array}{lll} 0 \le \lambda^x & \perp & Cx + d\lambda^x + eu \ge 0 \\ 0 \le \mu_{x,u} & \perp & \mu_{x,u} - 2(Cx+eu) \ge 0 \\ 0 \le \mu_1 & \perp & \mu_1 - 2\lambda_1^p \ge 0 \\ 0 \le \mu_2 & \perp & \mu_2 - 2\lambda_2^p \ge 0 \\ 0 \le \mu_p & \perp & \mu_p - 2p \ge 0 \\ 0 \le \lambda_1^{abs} & \perp & (Cx + eu + \lambda_{x,u}^{abs})\mathbb{1}_n \ge 0 \\ 0 \le \lambda_2^{abs} & \perp & (\lambda_{x,u}^{abs} - (Cx+eu))\mathbb{1}_n \ge 0 \\ & \lambda_1^{abs} + \lambda_2^{abs} = \lambda_p^{abs} & \\ & \lambda_1^p + \lambda_2^p = p & \\ & \lambda_{x,u}^{abs} = \mu_{x,u} - (Cx+eu) & \\ & \lambda_1^{abs} = \mu_1 - \lambda_1^p & \\ & \lambda_2^{abs} = \mu_2 - \lambda_2^p & \\ & \lambda_p^{abs} = \mu_p - p & \end{array}\right.$$

Noticing that we can use the two equalities on $\lambda_{x,u}^{abs}$ and λ_p^{abs} and insert them above, we have proven that λ^x and λ_1^p are equally defined by (11) or (12).■

Therefore, we infer that the right-hand side of the differential inclusion in (9) is equal to the right-hand side of system (10):
$$\dot{z} = \tilde{A}z + \tilde{B}\Lambda + \tilde{F}u, \tag{13}$$
where $\tilde{A}$, $\tilde{B}$, $\tilde{F}$ and Λ are easily identifiable from (10), and subject to the MLCP (12).

5. NUMERICAL OPTIMAL SOLUTION

5.1 Direct method

A first way to solve the optimal problem is to discretize directly the dynamics (4) and the cost (3) in order to obtain a constrained optimization problem. In fact, a simple discretization that we can use here is the following one:
$$\min_{u \in \mathbb{R}^N} \sum_{i=0}^{N} f^0(x_i, u_i),$$

$$s.t. \begin{cases} 0 \le \lambda_k^x & \perp Cx_{k+1} + d\lambda_k^x + eu_k \ge 0 \\ \dfrac{x_{k+1} - x_k}{h} & = \quad Ax_{k+1} + B\lambda_k^x + Fu_k \end{cases} \quad k = 0...N-1,$$

where the subscript k in z_k denotes the k-th step in the discretization of the variable $z(t)$ at time t_k, and h denotes the (here uniform) time-step. Moreover, we can choose to integrate the dynamics with an implicit (as presented here) or an explicit method. This is a Mathematical Program (MP) constrained by a Mixed Linear Complementarity Problem, written as MLCP$(G(\cdot,u,\cdot), H(\cdot,u,\cdot))$ with $G : \mathbb{R}^{n(N+1)} \times \mathbb{R}^{N+1} \times \mathbb{R}_+^N \to \mathbb{R}^{n(N+1)}$ and $H : \mathbb{R}^{n(N+1)} \times \mathbb{R}^{N+1} \times \mathbb{R}_+^N \to \mathbb{R}^N$, where

$$\begin{aligned} G_k(x,u,\lambda) &= \frac{x_{k+1} - x_k}{h} - Ax_{k+1} - B\lambda_k^x - Fu_k, \\ H_k(x,u,\lambda) &= Cx_{k+1} + d\lambda_k^x + eu_k, \end{aligned}$$

$k = 0,...,N-1$. We notice that we can isolate x_{k+1} in the dynamics, and reintroduce it in the complementarity conditions:

$$\begin{aligned} 0 \le \lambda_k^x \perp (d + hC(I-hA)^{-1}B)\lambda_k^x + \\ (e + hC(I-hA)^{-1}F)u_k + C(I-hA)^{-1}x_k \ge 0 \end{aligned}$$

As we can see here, this method is rather simple to implement, and can integrate further constraints. However, regarding dimensions, we see that a small discretization step $h > 0$ will increase dramatically the size of the system to be solved. Nonetheless, the solution of this problem can be the initial guess of the indirect method that we will present in the next section.

5.2 Indirect method

A second way to compute an approximate solution of the optimal control problem is to discretize (7) with (13)-(12). This method is known as the indirect method (since it is using an a priori study of the system, and the result obtained with the Pontryagin equations). Obviously, the choice of the discretization will eventually have an impact on the accuracy and the stability of the numerical solution. So as to have a first idea of the extent of these issues, these equations are discretized with an Euler scheme which will use implicit or explicit terms in its formulation, and we will investigate how these choices affect the solution. We present two formulations, that we name of *explicit* and *implicit* type. As we will see, these two formulations lead to different types of optimization problems.

Explicit type This first formulation leads to a problem where we can identify two (almost) independant problems at each step. Let us assume we already know variables values of x and p at time t_k, e.g. z_k, and we want to compute the solution at time t_{k+1}. We first solve the following discretization of (7) with (12):

$$\max_{v\in\mathbb{R}} \{\langle p_k, f^x(x_k, v)\rangle + p^0 f^0(x_k, v)\},$$

$$s.t. \begin{cases} 0 \le \lambda_k^x & \perp & Cx_k + d\lambda_k^x + ev \ge 0 \\ 0 \le \mu_{x,u} & \perp & \mu_{x,u} - 2(Cx_k + ev) \ge 0 \\ 0 \le \mu_p & \perp & \mu_p - 2p_k \ge 0 \\ 0 \le \lambda_1^{abs} & \perp & \mu_{x,u}\mathbb{1}_n \ge 0 \\ 0 \le \lambda_2^{abs} & \perp & (\mu_{x,u} - 2(Cx_k + ev))\mathbb{1}_n \ge 0 \\ 0 \le \mu_1 & \perp & \mu_1 - 2\lambda_{1k}^p \ge 0 \\ 0 \le \mu_2 & \perp & \mu_2 - 2\lambda_{2k}^p \ge 0 \\ & \lambda_1^{abs} + \lambda_2^{abs} = \mu_p - p_k & \\ & \lambda_{1k}^p + \lambda_{2k}^p = p_k & \\ & \mu_1 - \lambda_{1k}^p = \lambda_1^{abs} & \\ & \mu_2 - \lambda_{2k}^p = \lambda_2^{abs} & \end{cases}$$

Solving this problem will give us u_k and the associated Λ_k, which is unique for a given u_k as we seen from the derivation of system (12), except in the case where $Cx_k + eu_k = 0$. We can rewrite the complementarity conditions of this MPEC in the following compact form :

$$0 \le \Omega \perp \Delta\Omega + \Psi \ge 0$$

where $\Omega_k = (\lambda_k^x, \mu_{x,u}, \mu_p, \lambda_1^{abs}, \lambda_2^{abs}, \mu_1, \mu_2)^\intercal$,

$$\Delta = \begin{pmatrix} d & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbb{1}_n & 0 & 0 & 0 & 0 \\ 0 & \mathbb{1}_n & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbb{1}_n & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbb{1}_n & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbb{1}_n \end{pmatrix} \tag{14}$$

and Ψ is easily identifiable. Finally, we just need to integrate the dynamics (13), which is integrated with the following discretization:

$$\frac{z_{k+1} - z_k}{h} = \tilde{A}z_{k+1} + \tilde{B}\Lambda_k + \tilde{F}u_k.$$

Here again, we can use an implicit integration, or an explicit one by introducing z_k instead of z_{k+1} in the right-hand side.

Implicit type This second formulation is expressed in the form of a single Mathematical Program with Equilibrium Constraints (MPEC) solved at each timestep. Here, every variable will be used *implicitly*, as the dynamics is introduced inside the constraints of the MP. Namely, we need to solve at each step the following MPEC:

$$\max_{v\in\mathbb{R}} \{\langle p_{k+1}, f(x_{k+1}, v)\rangle + p^0 f^0(x_{k+1}, v)\}$$

$$s.t. \begin{cases} 0 \le \lambda_{k+1}^x & \perp & Cx_{k+1} + d\lambda_{k+1}^x + ev \ge 0 \\ 0 \le \mu_{x,u} & \perp & \mu_{x,u} - 2(Cx_{k+1} + ev) \ge 0 \\ 0 \le \mu_p & \perp & \mu_p - 2p_{k+1} \ge 0 \\ 0 \le \lambda_1^{abs} & \perp & \mu_{x,u}\mathbb{1}_n \ge 0 \\ 0 \le \lambda_2^{abs} & \perp & (\mu_{x,u} - 2(Cx_{k+1} + ev))\mathbb{1}_n \ge 0 \\ 0 \le \mu_1 & \perp & \mu_1 - 2\lambda_{1k+1}^p \ge 0 \\ 0 \le \mu_2 & \perp & \mu_2 - 2\lambda_{2k+1}^p \ge 0 \\ & \lambda_1^{abs} + \lambda_2^{abs} = \mu_p - p_{k+1} & \\ & \lambda_{1k+1}^p + \lambda_{2k+1}^p = p_{k+1} & \\ & \mu_1 - \lambda_{1k+1}^p = \lambda_1^{abs} & \\ & \mu_2 - \lambda_{2k+1}^p = \lambda_2^{abs} & \\ & \dfrac{z_{k+1} - z_k}{h} = \tilde{A}z_{k+1} + \tilde{B}\Lambda_{k+1} + \tilde{F}v & \end{cases}$$

Advantages and flaws of the indirect methods The main advantage of this kind of methods is that they produce usually highly accurate solutions (Rao, 2009). Also, in the framework developed here, a really small timestep will only increase the number of maximisation problems solved, but not the dimension of each of these, which is clearly an advantage over the direct method. However, the optimal control is computed as an open loop, and we have only necessary conditions but not sufficient ones, such that we still have to check that the solution that is found, is really optimal. On top of that, even though the systems are quite small, a maximisation problem will be solved at each step, which is a computationally hard problem.

6. 1D EXAMPLE

In order to illustrate the previous results, we now focus on a 1D example, written in the following generic form:

$$\begin{aligned} &\text{minimize } \int_0^T \left(x(t)^2 + u(t)^2\right) dt, \\ &\text{such that: } \begin{cases} \dot{x}(t) = ax(t) + b\lambda(t) + fu(t), \\ 0 \le \lambda(t) \perp d\lambda(t) + eu(t) \ge 0, \text{ a.e. on } [0,T] \\ x(0) = x_0,\ x(T) \text{ free,} \end{cases} \end{aligned} \tag{15}$$

where all variables are scalars, $d > 0$, $b, e \neq 0$.

6.1 Complete controllability conditions for this 1D example

In order to analyze this problem, we will first specify the necessary and sufficient complete controllability conditions in the 1D case. Applying theorem 1, we must check that the following system:

$$(a - \lambda)\zeta + c\eta = 0, \tag{16}$$

$$f\zeta + e\eta = 0, \tag{17}$$

$$\eta \ge 0, \tag{18}$$

$$b\zeta + d\eta \le 0, \tag{19}$$

has no solution $\lambda \in \mathbb{R}$ and $(\zeta, \eta) \neq 0$

If $e > 0$**:** we deduce through (17): $\eta = -\frac{f\zeta}{e}$.

(1) If $f = 0$, then $\eta = 0$. In (16), we can take $\lambda = a$. However, with (19), we have that $\zeta b \le 0$. Let us take $\zeta = -\text{sign}(b)$. Then we found a solution with $\zeta \neq 0$: the system is not completely controllable.

(2) If $f < 0$, then with (18), we have that $\zeta \ge 0$. Through (16), we take $\lambda = a + \frac{cf}{e}$.
- If $b \ge 0$, then (19) is a sum of positive terms which must be nonpositive, so $\eta = 0$ and $\zeta = 0$: the system is completely controllable.
- If $b < 0$, then (19) becomes $\zeta(b - \frac{fd}{e}) \le 0$ with $\zeta \ge 0$.
  - If $b - \frac{fd}{e} \le 0$ then we can take any $\zeta \ge 0$: the system is not completely controllable.
  - Otherwise, only $\zeta = 0$ suits, so $\eta = 0$, and then the system is completely controllable.

(3) If $f > 0$, then in (16), we take $\lambda = a + \frac{cf}{e}$. Through (18), we have that $\zeta \le 0$.
- If $b \le 0$, then (19) is a positive terms sum which must be nonpositive, so $\eta = 0$ and $\zeta = 0$: the system is completely controllable.
- If $b > 0$, then (19) becomes $\zeta(b - \frac{fd}{e}) \le 0$ with $\zeta \le 0$.
  - If $b - \frac{fd}{e} \ge 0$ then then we can take any $\zeta \le 0$: the system is not completely controllable.
  - Otherwise, only $\zeta = 0$ suits, so $\eta = 0$, and then the system is completely controllable.

If $e < 0$**:** we have the same cases as with $e > 0$ by inverting the sign of f.

6.2 Search for the explicit optimal solution

The dynamic system in (15) can be rewritten as:

$$\dot{x} = ax + fu + \frac{b}{d}\Pi_{\mathbb{R}_+}(-eu). \tag{20}$$

Therefore, the Hamiltonian function is written as:

$$H(x,p,u) = p\left(ax + fu + \frac{b}{d}\Pi_{\mathbb{R}_+}(-eu)\right) - \frac{1}{2}\left(x^2 + u^2\right). \tag{21}$$

Adjoint equation We notice that this equation is smooth in x. Therefore, using (6), the adjoint equation is smooth, and is written as:

$$\dot{p}(t) = -ap(t) + x(t).$$

We can even differentiate twice p, and obtain the following second-order differential equation:

$$\begin{aligned} \ddot{p} &= -a\dot{p} + \dot{x} \\ &= a^2p + fu + \frac{b}{d}\Pi_{\mathbb{R}_+}(-eu) \\ &= \begin{cases} a^2p + fu & \text{if } eu \ge 0 \\ a^2p + \left(f - \dfrac{be}{d}\right)u & \text{if } eu \le 0. \end{cases} \end{aligned}$$

Maximization of the Hamiltonian function We now search for an expression of the optimal control u^*, function of x and p, maximizing the Hamiltonian function $H(x,p,u^*)$. To that aim, we use the subdifferential of H with respect to u, written $\partial_u H(x,p,u)$, and the fact that if u^* maximizes H, then

$$0 \in \partial_u H(x,p,u^*).$$

In our problem, the subdifferential is written as

$$\partial_u H(x,p,u) = \begin{cases} \{fp - u\} & \text{if } eu > 0, \\ \left\{\left(f - \dfrac{eb}{d}\right)p - u\right\} & \text{if } eu < 0, \\ -\left[f - \dfrac{eb}{d}, f\right]p & \text{if } eu = 0. \end{cases}$$

We now only focus on the complete controllable cases in order to find a control u maximizing this function:

If $e > 0$**:** In that case, $\text{sgn}(eu) = \text{sgn}(u)$.

(1) We consider first $f < 0$.
- If $b > 0$, then if $p \le 0$, then $fp \ge 0$, $\left(f - \frac{eb}{d}\right)p \ge 0$, and if $p \ge 0$, then $fp \le 0$, $\left(f - \frac{eb}{d}\right)p \le 0$. We also notice that $0 \notin \left[f, f - \frac{eb}{d}\right]$. So we have:

$$u^* = \begin{cases} fp & \text{if } p \le 0, \\ \left(f - \dfrac{eb}{d}\right)p & \text{if } p \ge 0. \end{cases}$$

- If $b < 0$, then we must make sure that $f < \frac{eb}{d}$. We notice that in this case, $0 \notin \left[f, f - \frac{eb}{d}\right]$. We are then in the exact same case as the previous one, and therefore, the control is expressed the same way:

$$u^* = \begin{cases} fp & \text{if } p \leq 0, \\ \left(f - \dfrac{eb}{d}\right) p & \text{if } p \geq 0. \end{cases}$$

(2) We consider now $f > 0$.

- If $b < 0$, then if $p \leq 0$, then $fp \leq 0$, $\left(f - \frac{eb}{d}\right) p \leq 0$, and if $p \geq 0$, then $fp \geq 0$, $\left(f - \frac{eb}{d}\right) p \geq 0$. We also notice that $0 \notin \left[f, f - \frac{eb}{d}\right]$. So we have:

$$u^* = \begin{cases} fp & \text{if } p \geq 0, \\ \left(f - \dfrac{eb}{d}\right) p & \text{if } p \leq 0. \end{cases}$$

- If $b > 0$, then we must make sure that $f > \frac{eb}{d}$. We notice that in this case, $0 \notin \left[f, f - \frac{eb}{d}\right]$. We are then in the exact same case as the previous one, and therefore, the control is expressed the same way:

$$u^* = \begin{cases} fp & \text{if } p \geq 0, \\ \left(f - \dfrac{eb}{d}\right) p & \text{if } p \leq 0 \end{cases}.$$

If $e < 0$: we have the same cases as with $e > 0$ by inverting the sign of f.

Therefore, we can summarize this result as follows:

$$u^* = \begin{cases} fp & \text{if } efp \geq 0, \\ \left(f - \dfrac{eb}{d}\right) p & \text{if } efp \leq 0. \end{cases} \tag{22}$$

Final adjoint equation Finally, we use the optimal control found in (22) in the equation found on $\ddot{p}$. Surprisingly, we end up with a rather simple equation:

$$\ddot{p} = \begin{cases} \left(a^2 + f^2\right) p & \text{if } efp \geq 0, \\ \left(a^2 + \left(f - \dfrac{be}{d}\right)^2\right) p & \text{if } efp \leq 0, \end{cases}$$

which we rewrite in the more simple form:

$$\ddot{p} = \gamma(p)p \tag{23}$$

with $\gamma(p) > 0$ and piecewise constant.

Initial conditions We need now to find $p(0)$ such that $p(T) = 0$ (since $x(T)$ is free, according to the maximum principle). On top of that, we know that the initial value for the derivative $\dot{p}$ is given by:

$$\dot{p}(0) = x(0) - ap(0)$$

The phase portrait is depicted in Figure 1.

It is clear that, in order to have $p(T) = 0$, the sign of $p(0)$ is determined by the sign of the constants in the model:

- If $a > 0$, $x(0) > 0$, then $p(T) = 0 \implies p(0) < 0$,
- If $a > 0$, $x(0) < 0$, then $p(T) = 0 \implies p(0) > 0$,
- If $a < 0$, $x(0) > 0$, then $p(T) = 0 \implies p(0) < 0$,
- If $a < 0$, $x(0) < 0$, then $p(T) = 0 \implies p(0) > 0$.

We can summarize this by:

$$sgn(p(0)) = -sgn(x(0))$$

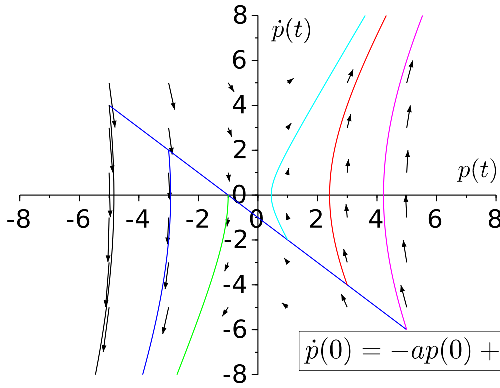

Fig. 1. Phase portrait of (23) - $a = 1$, $b = -0.5$, $d = 1$, $e = -2$, $f = 3$, $x(0) = -1$,

On top of that, p will always have the same sign on $[0, T]$, so the optimal control u^* given in equation (22) always has the same sign on $[0, T]$, and is smooth (since p is smooth). Furthermore, γ will be constant on $[0, T]$. Consequently, we know explicitly the solution $p(t)$ on $[0, T]$, namely:

$$p(t) = \frac{1}{2\sqrt{\gamma}} \left[\left((\sqrt{\gamma} - a)e^{\sqrt{\gamma}t} + (\sqrt{\gamma} + a)e^{-\sqrt{\gamma}t} \right) p(0) + \left(e^{\sqrt{\gamma}t} - e^{-\sqrt{\gamma}t} \right) x(0) \right].$$

In order to have $p(T) = 0$, we must take:

$$p(0) = -\frac{x(0)\left(e^{2\sqrt{\gamma}T} - 1\right)}{(\sqrt{\gamma} - a)e^{2\sqrt{\gamma}T} + \sqrt{\gamma} + a}.$$

From that, it is easy to have the expression of the optimal trajectory x, using the fact that

$$x(t) = \dot{p}(t) + ap(t).$$

6.3 Numerical results

Direct method The direct method gives rather good results, whatever the parameters in (15) used. We used the solver GAMS (available at `http://www.gams.com/`) which includes a powerful MPEC solver. The only trouble noticed was that some fluctuations around the analytical solution were found. Nonetheless, these fluctuations are still admissible in all the calculations we made. Some results are shown in Figure 2.

Indirect method The indirect method gives more disapointing results. As shown in Figure 3, even with the good initial value $p(0)$, this approach fails to give a good solution, close to the analytical one. Hope for a better solution with this method seems small since in general application, $p(0)$ should be found numerically. Here again, GAMS was used to solve the MPEC at each step. The reason why this is not working still is unknown to us: changing the parameters does not seem to enhance the precision of the numerical solution, nor the reduction of the timestep h. At each time step, the resolution of the MPEC seems to fail, the constraints being often largely violated. A first way to explain this may be found in matrix Δ introduced in (14) : even in this scalar example, it is of

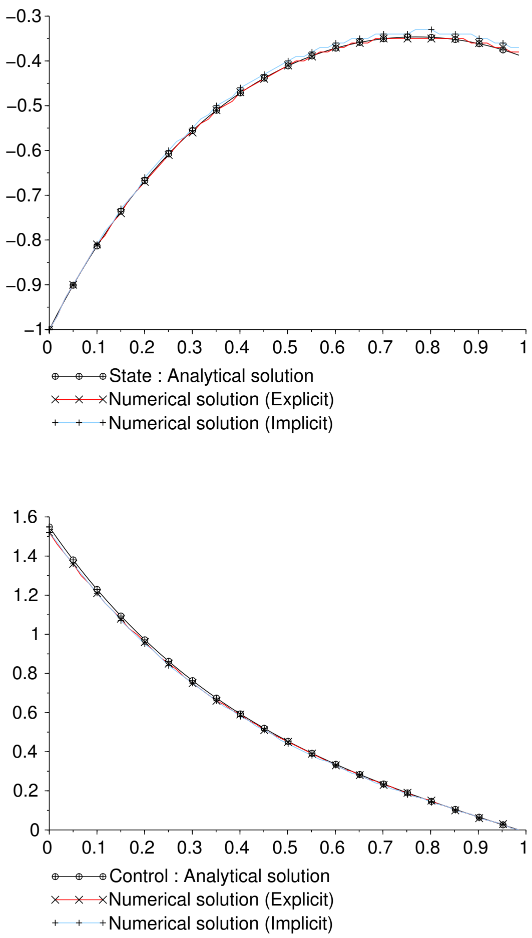

Fig. 2. Numerical solution: direct approach - $a = 1$, $b = -0.5$, $d = 1$, $e = -2$, $f = 3$, $x(0) = -1$, $N = 60$

rank 5, when Δ is 7×7. Further investigations are needed and are ongoing.

7. CONCLUSION

The results we found for the optimal control of linear complementarity systems are promising and we were able to formulate the first-order condition in a convenient framework using complementarity conditions. Even though we achieve getting these conditions under a suitable form, the numerical schemes derived from the indirect method are not satisfactory. The MPEC involved at each time-step seems computationnaly hard to solve, and needs further investigations. Some other possible solutions are currently under reasearch, as changing the mixed complementarity conditions involved, or using different algorithms for this maximisation.

ACKNOWLEDGEMENTS

This work has been partially supported by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01)

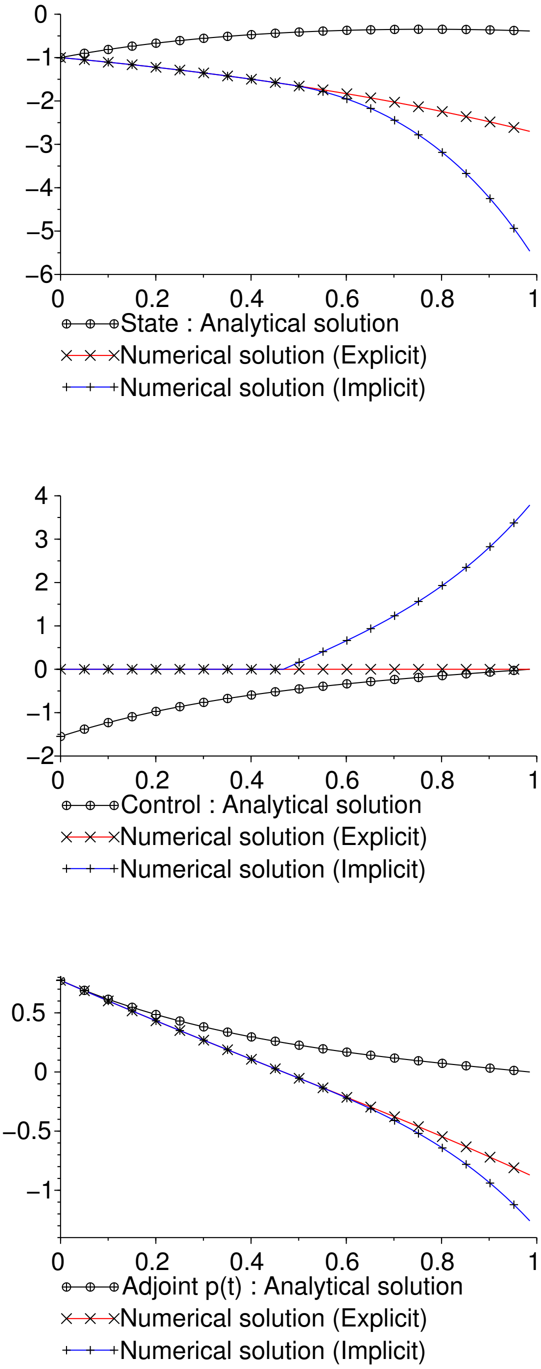

Fig. 3. Numerical solution: indirect approach - $a = 1$, $b = -0.5$, $d = 1$, $e = -2$, $f = 3$, $x(0) = -1$, $N = 60$

REFERENCES

Acary, V., Bonnefon, O., and Brogliato, B. (2011). *Nonsmooth Modeling and Simulation for Switched Circuits*, volume 69. Springer Science & Business Media.

Brogliato, B. (2003). Some perspectives on the analysis and control of complementarity systems. *IEEE Transactions on Automatic Control*, 48(6), 918–935.

Çamlibel, M. and Schumacher, J. (2001). On the Zeno behavior of linear complementarity systems. *Proceedings of the 40th IEEE Conference on Decision and Control*, 1, 346–351.

Çamlibel, M.K. (2007). Popov–Belevitch–Hautus type controllability tests for linear complementarity systems. *Systems & Control Letters*, 56(5), 381–387.

Çamlibel, M.K., Pang, J.S., and Shen, J. (2006). Conewise linear systems: non-zenoness and observability. *SIAM Journal on Control and Optimization*, 45(5), 1769–1800.

Cesari, L. (2012). *Optimization-Theory and Applications: Problems with Ordinary Differential Equations*, volume 17. Springer Science & Business Media.

Clarke, F. and De Pinho, M. (2010). Optimal control problems with mixed constraints. *SIAM Journal on Control and Optimization*, 48(7), 4500–4524.

Clarke, F.H. (1976). The maximum principle under minimal hypotheses. *SIAM Journal on Control and Optimization*, 14(6), 1078–1091.

Georgescu, C., Brogliato, B., and Acary, V. (2012). Switching, relay and complementarity systems: A tutorial on their well-posedness and relationships. *Physica D: Nonlinear Phenomena*, 241(22), 1985 – 2002.

Heemels, W.M.H., Camlibel, M.K., Schumacher, J.M., and Brogliato, B. (2011). Observer-based control of linear complementarity systems. *International Journal of Robust and Nonlinear Control*, 21(10), 1193–1218.

Pang, J.S. and Shen, J. (2007). Strongly regular differential variational systems. *IEEE Transactions on Automatic Control*, 52(2), 242–255.

Passenberg, B., Caines, P.E., Leibold, M., Stursberg, O., and Buss, M. (2013). Optimal control for hybrid systems with partitioned state space. *IEEE Transactions on Automatic Control*, 58(8), 2131–2136.

Rao, A.V. (2009). A survey of numerical methods for optimal control. *Advances in the Astronautical Sciences*, 135(1), 497–528.

Shen, J. (2014). Robust non-zenoness of piecewise affine systems with applications to linear complementarity systems. *SIAM Journal on Optimization*, 24(4), 2023–2056.